UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/056,546	01/24/2002	Sankar Basu	YOR920020018	8796

48813 7590 01/14/2009

LAW OFFICE OF IDO TUCHMAN (YOR)
ECM #72212
PO Box 4668
New York, NY 10163-4668

EXAMINER	
DEBROW, JAMES J	

ART UNIT	PAPER NUMBER
2176	

NOTIFICATION DATE	DELIVERY MODE
01/14/2009	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

pair@tuchmanlaw.com
idotuchman@gmail.com

Office Action Summary	Application No.	Applicant(s)
	10/056,546	BASU ET AL.
	Examiner	**Art Unit**
	JAMES J. DEBROW	2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on _08 October 2008_.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) _1,2,5-15,23 and 29-42_ is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) _1, 2, 5-15, 23 and 29-42_ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on _24 January 2002_ is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

This action is responsive to communications: Amendment filed 08 Oct 2008.

Claims 1, 2, 5-15, 23 and 29-42 are pending in this case. Claims 1, 33 and 41 are independent claims.

Applicant's Response

In Applicant's response dated 08 Oct 2008, Applicant amended claims 1, 6 and 13; canceled claim 24; added new claims 32-42; argued against all objections and rejection previously set forth in previous Office Action.

Specification

The specification is objected to as failing to provide proper antecedent basis for the claimed subject matter. See 37 CFR 1.75(d)(1) and MPEP § 608.01(o). Correction of the following is required: "*the confidence level* ***being inversely proportional to a distance of a <u>new feature</u> of the selected examples from a <u>separating hyperplane in an induced higher dimensional feature space</u>***," as recited in **Claims 1, 40 and 42**. The examiner realizes that this exact language is located in the Specification at Page 14, Lines 7-9; however, nothing in the Specification provides guidance in interpreting what this claim language means.

For example, exactly what is a “new feature,” and exactly what is a "separating hyperplane in an induced higher dimensional feature space?" Additionally, the examiner has the following questions after reading this claim language and the Specification:

1. Exactly how does a “new” feature differ from an “old” feature?
2. Exactly what is a “feature space?”
3. Exactly what makes a feature space “higher dimensional?”
4. Exactly how does an “induced” feature space differ from a “non-induced” feature space?
5. Exactly what is a “hyperplane” in a feature space?
6. Exactly how does a “separating” hyperplace differ from a “non-separating” hyperplane?

If the Specification is capable of answering any of the above questions, then Applicant must cite that portion of the Specification that answers the corresponding question and provide a rationale that explains in detail how the cited portion of the Specification answers the question.

Because the Specification fails to provide any guidance in interpreting the claim language and answering the above questions, the Specification does not provide clear support or antecedent basis for the terminology used in the claims, as required in Rule 75.

As an aside, the present invention appears to involve assigning keywords to a video, wherein the assigned keywords describe corresponding portions of the video and

thereby permit a user to "text search" the video. See Page 9, Line 28 through Page 10, Line 21 of the Specification and Figure 8 of the Drawings.

The specification is objected to as failing to provide proper antecedent basis for the claimed subject matter. See 37 CFR 1.75(d)(1) and MPEP § 608.01(o). Correction of the following is required: the "*information measures*" of **Claim 5**. The Specification fails to provide any guidance in interpreting what this claim language means. Thus, the Specification does not provide clear support or antecedent basis for the terminology used in the claims, as required in Rule 75.

The specification is objected to as failing to provide proper antecedent basis for the claimed subject matter. See 37 CFR 1.75(d)(1) and MPEP § 608.01(o). Correction of the following is required: the "*cepstral coefficient*" of **Claim 37**. The Specification fails to provide any guidance in interpreting what this claim language means. Thus, the Specification does not provide clear support or antecedent basis for the terminology used in the claims, as required in Rule 75.

These objections may be obviated by cancelling the claims or cancelling the unsupported language in the claims.

Drawings

The drawings are objected to because the text in the GUIs of Figures 8-13 is small, unfocused and difficult/impossible to read. Applicant should amend Figures 8-13 so that all wording is easily readable.

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as “amended.” If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either “Replacement Sheet” or “New Sheet” pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Claim Objections

Claim 6 is objected to because of the following informalities: Claims 6 contains an incorrect status identifier. The claim status identifier should be "currently amended", as the claim has been amended. See MPEP 1.121(c). Appropriate correction is required.

Claim Rejections - 35 USC § 112

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claim 32 is rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter, "*surveillance data*", which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

Claim 37 is rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter, *"cepstral coefficient"*, which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention. The specification only recites the term "cepstral coefficient" in

section 0029, which does not provide sufficient information to it's meaning and use in this context.

Claim 38 is rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter, *"zero crossing"*, which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention. The specification only recites the term "zero crossing" in section 0029, which does not provide sufficient information to it's meaning and use in this context.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1, 2, 6, 10 and 32 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler et al. (Patent No.: US 6,804,684, B2; Filed May 7, 2001) (hereinafter "Stubler") in view of Toyama (Patent No.: US 6,816,847 B1; Filling Date: Sep. 23, 1999).

In regards to independent claim 1 Stubler discloses a *method for generating persistent annotations of multimedia content, comprising one or more repetitions of the following steps:*

actively selecting examples of multimedia content to be annotated by a user (fig. 2, fig. 6-7, col. 3 line 46 – col. 4 line 12, col. 8 lines 18-23, and col. 9 line 65 – col. 10 line 18; col. 11, lines 3-35; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously. Stubler also disclose automatically generating one or more captions or labels for the acquired image automatically without any user intervention.).

accepting input annotations from said user for said selected examples (fig. 2, col. 3 line 46 – col. 4 line 12, and col. 8 lines 18-55; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously.).

propagating said input annotations to other instances of multimedia content (fig. 2; col. 3 line 46 – col. 4 line 12; col. 8 lines 18-55; col. 2 line 59 – col. 3 line 10; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously.).

storing said input annotations and said propagated annotations (col. 8 lines 18-55; col. 9, line 65-col. 10, line 18; Stubler discloses storing input annotations and said propagated annotations.).

Stubler does not expressly disclose *wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected,*

wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space.

Toyama teaches *wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected* (col. 5, lines 15-65; Toyama teaches a set of images in the training data that are selected and assigned an aesthetic score, by a group of professionals, based on how aesthetically pleasing or how aesthetically poor *(most ambiguous)* the image may be based on whatever criteria the person wishes to use for deeming the aesthetics of the image, or according to some standard specified by the survey. Each image may, for example, be scored, for example, on a classification basis, where there are a number of categories, such as excellent, good, average, or poor. Toyama also teaches the invention is not limited to a particular number or a particular type of image features used by the classifier to discern commonality among like-judged images and their corresponding aesthetic scores. Using the broadest reasonable interpretation, the Examiner concludes Toyama invention to

also include discerning most ambiguous images and their corresponding aesthetic scores.).

wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space (col. 5, line 15-col. 7, line 19; Toyama teaches a set of images in the training data that are selected and assigned an aesthetic score, by a group of professionals. Each image may, for example, be scored, for example, on a classification basis, where there are a number of categories/classificatory class, such as excellent, good, average, or poor. Toyama also teaches Support Vector Machines builds classifiers by identifying a hyperplane that separates a set of positive and negative example with a maximum margin. The margin is defined by the distance of the hyperplane to the nearest positive and negative cases of each class. Further, applicant disclosed within the specification (0008), for SVM classifiers the distance of an unlabeled data-point from the separating hyperplane in the high dimensional feature space could be taken as a measure of uncertainty (alternatively, a measure of confidence in classification) of the data-point. Therefore, using the broadest reasonable interpretation, the Examiner concludes that positive cases/example would receive a high/excellent aesthetic score (confidence level) and negative cases/example would receive a low/poor aesthetic score. Also as previously stated, the greater the margin/distance between the positive cases and the negative, the greater margin/distance of the hyperplane separating the cases. Therefore, the most negative

case/example with the lowest/poorest aesthetic score (confidence level) would also have the greatest margin/distance of separation of the hyperplane from the most positive case/example. Thus the aesthetic score (confidence level) is inversely proportional to the margin/distance of separation of the hyperplane in an induced higher dimensional feature space.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Toyama with Stubler for the benefit of providing and correlating the score for images based on features of the image to make up a training set (col. 1, lines 57-66).

In regards to dependent claim 2, Stubler discloses *wherein the step of actively selecting is performed using a selection technique selected from deterministic* (col. 4 line 64 – col. 5 line 19, Stubler discloses using image similarity based upon object measure or determinable metadata to provide a mechanism for assigning captions or semantic labels to multiple images simultaneously.).

Stubler in view of does not expressly disclose *actively selecting is performed using a selection technique of probabilistic.*

However Toyama teaches *actively selecting is performed using a selection technique of probabilistic* (col. 7, lines 7-19; Toyama teaches the classifier generates probabilities that the image falls into one or more aesthetical classes, or just the aesthetical class into which the image has the highest probabilities to be located.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Toyama with Stubler for the benefit of providing and correlating the score for images based on features of the image to make up a training set (col. 1, lines 57-66).

In regards to dependent claim 6, Stubler discloses *wherein the multimedia content comprises one or more types selected from the group consisting of: audio, video, Web pages, time series data, sensor data, and XML data* (col. 6. line 31-col. 7, line 30; col. 8, lines 18-39; col. 9, lines 26-39; Stubler discloses well known forms of capture metadata including time/date (time series data) and global positioning (GPS) data provided by the capture device.).

In regards to dependent claim 10, Stubler discloses *the method of claim 1, wherein the input annotations are created by means of steps selected from the group consisting of: creating new annotations, deleting existing annotations, rejecting proposed annotations, and modifying annotations* (col. 9, line 65-col. 10, line 18; Stubler disclose an interactive user verification stage in which the user may select and/or edit captions and label.).

In regards to dependent claim 32, Stubler discloses *the method of claim 1, wherein the multimedia content comprises surveillance data* (col. 6. line 31-col. 7, line 30; col. 8, lines 18-39; col. 9, lines 26-39; Stubler discloses well known forms of capture

metadata including time/date and global positioning (GPS) data (surveillance data) provided by the capture device.).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler in view of Toyama, further in view of Lennon et al. (Patent No.: 6,718,063 B1; Filed Dec. 10, 1999) (hereinafter "Lennon").

In regards to dependent claim 5, Stubler in view of Toyama does not expressly disclose *wherein an optimization criterion for active selection includes one or more criteria selected from the group consisting of: information measures.*

However Lennon teaches *wherein an optimization criterion for active selection includes one or more criteria selected from the group consisting of: information measures* (col. 3, lines 34-36; col. 3, line 63-col. 4, line 4; Lennon teaches comparing the distance metric with a predetermined threshold in order to determine the similarity of

the images. Thus Lennon teaches active selection includes one or more criteria selected from the group consisting of: information measures, using distance and threshold measures.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lennon with Stubler in view of Toyama for the benefit of using a probabilistic method, which results in each assigned label for a region having an associated probability or likelihood of the label being correctly assigned (col. 3, lines 23-27).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Claims 7, 8, 11, 13-15, 29-31, 33-36 and 39-42 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler in view of Toyama further in view of Lipson et al. (Patent No.: 5,963,670; Filed: Feb. 12, 1996) (hereinafter "Lipson").

In regards to dependent claim 7, Stubler in view of Toyama does not expressly disclose *the method of claim 1, wherein the input annotations are created by a user with reference to a vocabulary.*

However Lipson teaches *the method of claim l, wherein the input annotations are created by a user with reference to a vocabulary* (col. 9, lines 23-35; Lipson teaches a rich vocabulary to differentiate between many classes of images.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 8, Stubler in view of Toyama does not expressly disclose *the method of claim 7, wherein the vocabulary contains one or more items selected from the group consisting of: terms, concepts, labels, and annotations.*

However Lipson teaches *the method of claim 7, wherein the vocabulary contains one or more items selected from the group consisting of: terms, concepts, labels, and annotations* (col. 9, lines 23-42; Lipson teaches a rich vocabulary to differentiate between many classes of images. It has been established and is commonly known that vocabularies typically contain *terms and concepts.*).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 11, Stubler in view of Toyama does not expressly disclose the method of claim 7, wherein the vocabulary is adaptively or dynamically organized and/or limited by the system or the user.

However Lipson teaches *the method of claim 7, wherein the vocabulary is adaptively or dynamically organized and/or limited by the system or the user* (col. 9, lines 23-35.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 13, Stubler discloses *the method of claim 1, wherein the propagation of annotations is determined deterministically or probabilistically* (col. 4 line 64 – col. 5 line 19; Stubler discloses using image similarity based upon object measure or determinable metadata to provide a mechanism for assigning captions or semantic labels to multiple images simultaneously.).

Stubler in view of Toyama does not expressly disclose *the use of models for each annotation or for joint annotations.*

However Lipson teaches *the use of models for each annotation or for joint annotations* (col. 2, lines 56-59; Lipson teaches an image model for classifying or detecting images.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 14, Stubler in view of Toyama does not expressly disclose *the method of claim 13, wherein the models are created or learned automatically or semi-automatically and/or are updated adaptively from interaction with the user.*

However Lipson teaches *wherein the models are created or learned automatically or semi-automatically and/or are updated adaptively from interaction with the user* (col. 9, line 37-col. 10, line 22; Lipson teaches a method for generating a class model.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 15, Stubler in view of Toyama does not expressly disclose *the method of claim 13, wherein the models are based on nearest neighbor voting or variants, parametric or statistical models, expert systems, rule-based systems, or hybrid techniques.*

However Lipson teaches *wherein the models are based on nearest neighbor voting or variants* (col. 2, lines 56-59; col. 13, line 2-col. 14, line18).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 29, Stubler in view of Toyama does not expressly disclose *the method of claim 13, wherein the models are based on expert systems.*

However Lipson teaches *wherein the models are based on expert systems* (col. 8, lines 53-60; Lipson teaches a model in which a user has personal knowledge of relevant relative relationship for a given class of images. Thus Lipson teaches wherein the models are based on expert systems.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 30, Stubler in view of Toyama does not expressly disclose *the method of claim 13, wherein the models are based on rule-based systems.*

However Lipson teaches *wherein the models are based on rule-based systems* (col. 3, lines 29-35; col. 17, lines 10-42; Lipson teaches deformable image template models. Is has been established and is well known in the art that template are rule-based.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 31, Stubler in view of Toyama does not expressly disclose *the method of claim 13, wherein the models are based on hybrid techniques.*

However Lipson teaches *wherein the models are based on hybrid techniques* (col. 2, lines 42-54; col. 13, line 56-col. 14, line 18; Lipson teaches a model which is provided from spatial and photometric properties of an image. Thus Lipson teaches models are based on hybrid techniques, a combination of models/relationships.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to independent claim 33, Stubler discloses a *method for generating persistent annotations of multimedia content, comprising one or more repetitions of the following steps:*

actively selecting examples of multimedia content to be annotated by a user (fig. 2, fig. 6-7, col. 3 line 46 – col. 4 line 12, col. 8 lines 18-23, and col. 9 line 65 – col. 10 line 18; col. 11, lines 3-35; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously. Stubler also disclose automatically generating one or more captions or labels for the acquired image automatically without any user intervention.).

accepting input annotations from said user for said selected examples (fig. 2, col. 3 line 46 – col. 4 line 12, and col. 8 lines 18-55; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously.).

propagating said input annotations to other instances of multimedia content (fig. 2; col. 3 line 46 – col. 4 line 12; col. 8 lines 18-55; col. 2 line 59 – col. 3 line 10; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously.).

storing said input annotations and said propagated annotations (col. 8 lines 18-55; col. 9, line 65-col. 10, line 18; Stubler discloses storing input annotations and said propagated annotations.).

using said input annotations as training data to update the model.

Stubler does not expressly disclose *wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected,*

wherein the at least one criterion is measured according to a model.

using said input annotations as training data to update the model.

Toyama teaches *wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected* (col. 5, lines 15-65; Toyama teaches a set of images in the training data that are selected and assigned an aesthetic score, by a group of professionals, based on how aesthetically pleasing or how aesthetically poor *(most ambiguous)* the image may be based on whatever criteria the person wishes to use for deeming the aesthetics of the image, or according to some standard specified by the survey. Each image may, for example, be scored, for example, on a classification basis, where there are a number of categories, such as excellent, good, average, or poor. Toyama also teaches the invention is not limited to a particular number or a particular type of image features used by the classifier to discern commonality among like-judged images and their corresponding aesthetic scores. Using the broadest reasonable interpretation, the Examiner concludes Toyama invention to also include discerning most ambiguous images and their corresponding aesthetic scores.).

using said input annotations as training data (col. 1, lines 57-60; col. 5, lines 30-46; Toyama teaches a set of images, in which each has been assigned an aesthetic score, in the training set.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Toyama with Stubler for the benefit of providing and correlating the score for images based on features of the image to make up a training set (col. 1, lines 57-66).

Stubler in view of Toyama does not expressly disclose *wherein the at least one criterion is measured according to a model.*

data to update the model.

However Lipson teaches *wherein the at least one criterion is measured according to a model* (col. 2, lines 19-55; col. 13, lines 38-55; col. 15, lines 28-33; Lipson teaches models may be selected in accordance with a variety of factors but not limited to the class of images to which the model is to be applied. Lipson also teaches a user may specify images which should or should not be described by the class model. Thus Lipson teaches wherein the at least one criterion is measured according to a model.).

data to update the model (col. 13 lines 2-7; Lipson teaches a user updating features of a model.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 34, Stubler discloses *the method of claim 33, further comprising repeating the step of selecting examples of multimedia content* (fig. 2, fig. 6-7, col. 3 line 46 – col. 4 line 12, col. 8 lines 18-23, and col. 9 line 65 – col. 10 line 18; col. 11, lines 3-35; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously. Stubler also disclose automatically generating one or more captions or labels for the acquired image automatically without any user intervention.).

Stubler does not expressly disclose *wherein the repeated step of selecting examples of multimedia content is performed according to the model updated by using said input annotations as training data.*

using said input annotations as training data (col. 1, lines 57-60; col. 5, lines 30-46; Toyama teaches a set of images, in which each has been assigned an aesthetic score, in the training set.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Toyama with Stubler for the benefit of providing and

correlating the score for images based on features of the image to make up a training set (col. 1, lines 57-66).

Stubler in view of Toyama does not expressly disclose *selecting examples of multimedia content is performed according to the model updated.*

However Lipson teaches *selecting examples of multimedia content is performed according to the model updated* (col. 2, lines 19-55; col. 13 lines 2-7; col. 13, lines 38-55; col. 15, lines 28-33; Lipson teaches models may be selected in accordance with a variety of factors but not limited to the class of images to which the model is to be applied. Lipson also teaches a user may specify images which should or should not be described by the class model. Lipson further teaches a user updating features of a model. Thus Lipson teaches selecting examples of multimedia content is performed according to the model updated.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 35, Stubler in view of Toyama does not expressly disclose *the method of claim 33, wherein the model uses at least one feature representation.*

However Lipson teaches *wherein the model uses at least one feature representation* (col. 2, lines 19-55; col. 13 lines 25-37; Lipson teaches image models for classifying and detecting images includes a plurality of image patches couple by a plurality of relative image patch property relationships. For example the patch properities may correspond to properties including but not limited to one or more of spatial, color, photometric, texture, luminance and shape properties of the image pateches.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 36, Stubler in view of Toyama does not expressly disclose *the method of claim 35, wherein the at least one feature representation comprises a texture.*

However Lipson teaches *wherein the at least one feature representation comprises a texture* (col. 2, lines 19-55; col. 13 lines 25-37; Lipson teaches image models for classifying and detecting images includes a plurality of image patches couple by a plurality of relative image patch property relationships. For example the patch properities may correspond to properties including but not limited to one or more of spatial, color, photometric, texture, luminance and shape properties of the image pateches.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 39, Stubler disclose *the method of claim 33, further comprising performing user verification when the step of propagating said input annotations has been performed with least confidence* (col. 9, line 65-col. 10, line 18; Stubler disclose an interactive user verification stage in which the user may select and/or edit captions and label.).

In regards to dependent claim 40, Stubler does not expressly disclose *the method of claim 33, wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space.*

Toyama teaches *wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space* (col. 5, line 15-col. 7, line 19; Toyama teaches a set of images in the training data that are selected and assigned an aesthetic score, by a group of professionals. Each image may, for example, be scored, for example, on a

classification basis, where there are a number of categories/classificatory class, such as excellent, good, average, or poor. Toyama also teaches Support Vector Machines builds classifiers by identifying a hyperplane that separates a set of positive and negative example with a maximum margin. The margin is defined by the distance of the hyperplane to the nearest positive and negative cases of each class. Further, applicant disclosed within the specification (0008), for SVM classifiers the distance of an unlabeled data-point from the separating hyperplane in the high dimensional feature space could be taken as a measure of uncertainty (alternatively, a measure of confidence in classification) of the data-point. Therefore, using the broadest reasonable interpretation, the Examiner concludes that positive cases/example would receive a high/excellent aesthetic score (confidence level) and negative cases/example would receive a low/poor aesthetic score. Also as previously stated, the greater the margin/distance between the positive cases and the negative, the greater margin/distance of the hyperplane separating the cases. Therefore, the most negative case/example with the lowest/poorest aesthetic score (confidence level) would also have the greatest margin/distance of separation of the hyperplane from the most positive case/example. Thus the aesthetic score (confidence level) is inversely proportional to the margin/distance of separation of the hyperplane in an induced higher dimensional feature space.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Toyama with Stubler for the benefit of providing and

correlating the score for images based on features of the image to make up a training set (col. 1, lines 57-66).

In regards to independent claim 41, Stubler discloses a *method for generating persistent annotations of multimedia content, comprising one or more repetitions of the following steps:*

actively selecting examples of multimedia content to be annotated by a user (fig. 2, fig. 6-7, col. 3 line 46 – col. 4 line 12, col. 8 lines 18-23, and col. 9 line 65 – col. 10 line 18; col. 11, lines 3-35; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously. Stubler also disclose automatically generating one or more captions or labels for the acquired image automatically without any user intervention.).

accepting input annotations from said user for said selected examples (fig. 2, col. 3 line 46 – col. 4 line 12, and col. 8 lines 18-55; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously.).

propagating said input annotations to other instances of multimedia content (fig. 2; col. 3 line 46 – col. 4 line 12; col. 8 lines 18-55; col. 2 line 59 – col. 3 line 10; Stubler discloses unlabeled image regions being presented to the user so that the user can apply a caption or label to all of the regions simultaneously.). *and*

storing said input annotations and said propagated annotations (col. 8 lines 18-55; col. 9, line 65-col. 10, line 18; Stubler discloses storing input annotations and said

propagated annotations.).

Stubler does not expressly disclose *wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected;*

wherein a rectangular region of an image is associated with at least one of said input annotations;

Toyama teaches *wherein the examples of multimedia content are selected based on at least one criterion for achieving a maximal disambiguation result such that only those examples which are most ambiguous are selected* (col. 5, lines 15-65; Toyama teaches a set of images in the training data that are selected and assigned an aesthetic score, by a group of professionals, based on how aesthetically pleasing or how aesthetically poor *(most ambiguous)* the image may be based on whatever criteria the person wishes to use for deeming the aesthetics of the image, or according to some standard specified by the survey. Each image may, for example, be scored, for example, on a classification basis, where there are a number of categories, such as excellent, good, average, or poor. Toyama also teaches the invention is not limited to a particular number or a particular type of image features used by the classifier to discern commonality among like-judged images and their corresponding aesthetic scores. Using the broadest reasonable interpretation, the Examiner concludes Toyama invention to

also include discerning most ambiguous images and their corresponding aesthetic scores.).

Stubler in view of Toyama does not expressly disclose *wherein a rectangular region of an image is associated with at least one of said input annotations*

However Lipson teaches *wherein a rectangular region of an image is associated with at least one of said input annotations* (col. 10, line 5-col. 11, line 17; col. 14, line 44-col. 15, line 27; col. 17, lines 47-61; col. 21, lines 3-34; Figs. 4-10; Lipson teaches a plurality of image regions (rectangular region) which may be classified (annotations) based on image characteristics and properties.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Lipson with Stubler in view of Toyama for the benefit of providing a class model, which can be used to detect images of that class in a database (col. 2, lines 39-41).

In regards to dependent claim 42, Stubler does not expressly disclose *the method of claim 41, wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space.*

Toyama teaches *wherein the at least one criterion includes a confidence level of the selected examples, the confidence level being inversely proportional to a distance of a new feature of the selected examples from a separating hyperplane in an induced higher dimensional feature space* (col. 5, line 15-col. 7, line 19; Toyama teaches a set of images in the training data that are selected and assigned an aesthetic score, by a group of professionals. Each image may, for example, be scored, for example, on a classification basis, where there are a number of categories/classificatory class, such as excellent, good, average, or poor. Toyama also teaches Support Vector Machines builds classifiers by identifying a hyperplane that separates a set of positive and negative example with a maximum margin. The margin is defined by the distance of the hyperplane to the nearest positive and negative cases of each class. Further, applicant disclosed within the specification (0008), for SVM classifiers the distance of an unlabeled data-point from the separating hyperplane in the high dimensional feature space could be taken as a measure of uncertainty (alternatively, a measure of confidence in classification) of the data-point. Therefore, using the broadest reasonable interpretation, the Examiner concludes that positive cases/example would receive a high/excellent aesthetic score (confidence level) and negative cases/example would receive a low/poor aesthetic score. Also as previously stated, the greater the margin/distance between the positive cases and the negative, the greater margin/distance of the hyperplane separating the cases. Therefore, the most negative case/example with the lowest/poorest aesthetic score (confidence level) would also have the greatest margin/distance of separation of the hyperplane from the most

positive case/example. Thus the aesthetic score (confidence level) is inversely proportional to the margin/distance of separation of the hyperplane in an induced higher dimensional feature space.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Toyama with Stubler for the benefit of providing and correlating the score for images based on features of the image to make up a training set (col. 1, lines 57-66).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Claims 9 and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler in view of Toyama, further in view of Chino et al. (Patent No.: 6,118,888; Filed: Feb. 25, 1998) (hereinafter “Chino”).

In regards to dependent claim 9, Stubler in view of Toyama does not expressly disclose *the method of claim 1, wherein the process of creating input annotations by the user involves multimodal interaction with the user.*

However Chino teaches *wherein the process of creating input annotations by the user involves multimodal interaction with the user* (col. 3, lines 17-65; col. 30, lines 13-15; Chino teaches creating input annotations by the user involves multimodal interaction with the user.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Chino with Stubler in view of Toyama for the benefit of providing a multimodal interface apparatus and method to smoothly communicate between the user and the apparatus using the user’s gaze object (col. 3, lines 17-20).

In regards to dependent claim 12, Stubler in view of Toyama does not expressly disclose *the method of claim 9, wherein the multimodal interaction involves one or more elements selected from the group consisting of: gaze detection, finger pointing, expression detection, and effective computing methods for sensing a user's state.*

However Chino teaches *wherein the multimodal interaction involves one or more*

elements selected from the group consisting of: gaze detection, finger pointing, expression detection, and effective computing methods for sensing a user's state (Abstract; col. 17, lines 14-58; Chino teaches a multimodal interface which obtains position information of the objected pointed to by the gesture input, the user’s face or hand moved by the gesture input.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Chino with Stubler in view of Toyama for the benefit of providing a multimodal interface apparatus and method to smoothly communicate between the user and the apparatus using the user’s gaze object (col. 3, lines 17-20).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Claim 23 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler in view of Toyama further in view of Neal et al. (Patent No.: US 6,697,799 B1; Effective Filling Date: Sep. 10, 1999) (hereinafter "Neal").

In regards to dependent claim 23, Stubler in view of Toyama does not expressly disclose *wherein the at least one criterion includes an ambiguity level of the selected examples.*

However Neal teaches *wherein the at least one criterion includes an ambiguity level of the selected examples* (col. 11, lines 1-47; Neal teaches a classification confidence score which determines the level of confidence in which a category is likely to be correct during classification. If the item has a high confidence, then it can be classified directly. If the confidence level is low, then the results can be sent to the user interface for review and selection by the operator. Using the broadest reasonable interpretation, the Examiner has determined that the "confidence level" as taught by Neal is analogous with the "ambiguity level" of the current invention.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Neal with Stubler in view of Toyama for the benefit of comparing confidence scores of items for all selected classifications and classifying the items based on confidence score comparison (col. 2, lines 36-42).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Claim 37 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler and Toyama in view of Lipson, further in view Tourneret et al, (NPL: "Study of the Cepstral Coefficient Probability Density Function"; Published 1992, IEEE), (hereinafter "Tourneret").

In regards to dependent claim 37, Stubler does not expressly disclose *the method of claim 35, wherein the at least one feature representation comprises a cepstral coefficient.*

Toyama teaches coefficients of various image transformations of image regions (col. 5, lines 47-65).

However Stubler and Toyama in view of Lipson does not expressly disclose *wherein the at least one feature representation comprises a cepstral coefficient*.

Tourneret teaches *a cepstral coefficient* (Tourneret teaches using cepstral coefficient for pattern recognition and classification. Tourneret also teaches that *cepstral coefficient* it has been established and well known in the art. (pages 440-443).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Tourneret with Stubler and Toyama in view of Lipson for the benefit of providing a recursive method for computing the estimated cepstral coefficient probability density (5-Conclusion, page 443).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Claim 38 is rejected under 35 U.S.C. 103(a) as being unpatentable over Stubler and Toyama in view of Lipson, further in view Marr et al, (NPL: "Theory of Edge Detection"; Published 1980), (hereinafter "Marr").

In regards to dependent claim 38, Stubler and Toyama in view of Lipson does not expressly disclose *the method of claim 35, wherein the at least one feature representation comprises zero crossings.*

Marr teaches *wherein the at least one feature representation comprises zero crossings* (page 195, line 1-3 & 11-13; Marr teaches a zero-crossing segment in a Gaussian filtered image consists of a linear segment L of zero-crossings in the second directional derivative operator whose direction lies perpendicular to L. The set of zero-crossing segments together with their amplitudes, constitutes a primitive symbolic representation of the changes taking place within one region of the spectrum of an image. As described within the specification of the current invention, the gaussian mixture model classifiers may use the zero crossing feature to measures of degree of uncertainty of classification (0007, 0028-0029. Thus the Gaussian filtered image as taught in Marr teaches least one feature representation comprises zero crossings.).

Therefore at the time of the invention it would have been obvious to one of ordinary skill in the art to combine Marr with Stubler and Toyama in view of Lipson for the benefit of providing a theory of edge detection, which teaches the operation of forming oriented zero-crossing segments on the image (1st para, page 197).

Note

It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See MPEP 2123.

Response to Arguments

Previous claim objection is withdrawn.

Applicant argues, *"while the specification states that the distance of an unlabeled data-point from the separating hyperplane in the high dimensional feature space could be taken as a measure of uncertainty, the specification does not state that this distance can only be employed as a measure of uncertainty. It is noted that the language, "could be", clearly expresses a possibility rather than an absolute requirement. Therefore, the Applicants respectfully submit that it is not inherent that Toyama employs a distance of an unlabeled data-point from a separating hyperplane as a measure of uncertainty."* (Remarks page 16).

The Examiner disagrees.

The specification (008) of the current invention expressly states "For SVM (Support Vector Machine) classifiers the distance of an unlabeled data-point from the separating hyperplane in the high dimensional feature space could be taken as a measure of uncertainty of the data-point."

Likewise, Toyama teaches a SVM that which in employed in defining the margin as a distance of the hyperplane to the nearest positive and negative cases for each class. Thus the Examiner concludes the margin as taught in Toyama is analogous to

the measure of uncertainty of the data-point as determined by the SVM as describes within the claimed invention (col. 6, lines 19-31).

Further, in response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., *"a distance of an unlabeled data-point from a separating hyperplane as a measure of uncertainty"*) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Additional Applicant's arguments have been considered but are moot in view of the new ground(s) of rejection. A new ground(s) of rejection is made in view of Stubler, Lennon, Lipson, Neal, Toyama, Chino, Tourneret and/or Marr.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to James J. Debrow whose telephone number is 571-272-5768. The examiner can normally be reached on 8:00-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doug Hutton can be reached on 571-272-4137. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

JAMES DEBROW
EXAMINER
ART UNIT 2176

/Doug Hutton/
Doug Hutton
Supervisory Primary Examiner
Technology Center 2100